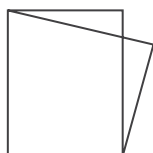


Den skønnede evidens

Sanne Haase,
postdoc, ph.d.,
Center for
Forskningsanalyse,
Aarhus Universitet



Fortællingen om Danmark som "vidensnation" er fremtrædende i vores samtid og indebærer en tiltro til, at vi skal konkurrere på vores evne til at udtænke smartere løsninger end andre lande. Evidens-basering ses ofte som nøglen til at identificere, hvilke løsninger der er "smarte". Evidensbevægelsen har bredt sig langt ind i den professionelle praksis med massiv kritik i sit kølvand. I stedet for helt af afvise et evidensbegreb, der rettelig ofte er for insensitivt over for en dynamisk og kompleks virkelighed, er det intentionen med denne artikel at forsøge at omdefinere og udvide evidensbegrebet, så det i højere grad rummer de professionelles egne vidensformer og deres professionelle skøn.

Introduktion

Når offentlige ydelser skal udføres, er det i alles interesse, at det gøres på den bedst mulige måde, så indsatserne bliver så virkningsfulde som overhovedet muligt. Allerede ved initieringen er der således en forventning om en intenderet effekt af at iværksætte politiske initiativer og interventioner i den professionelle praksis.¹

Udviklingen af en samfundskonstruktion, hvor viden skal informere den politiske og den sociale praksis, er ikke nogen ny foreteelse, men kan spores tilbage til fx Machiavelli (Hammersley, 2013). De senere årtiers "kamp" om evidensen (Krogstrup, 2011) har derimod handlet om, hvad viden af typen evidens skal forstås som, hvem der har legitimt krav på at have en sådan viden, og i særdeleshed, hvilke valide metoder der kan tilvejebringe en sådan evidens.

Det dominerende evidensbegreb knytter sig til forestillingen om, at der på forhånd er "bevis" for, at fx en indsats virker efter hensigten. Den valgte indsats skal være den bedste relativt til andre.

Evidensbegrebet er ikke forbeholdt den offentlige sektor, men synes at være særligt udbredt her. For ti år siden afdækkede Rieper og Hansen (2007), hvordan begrebet syntes at have bevæget sig fra en forskningsintern metodedebat, der var særligt relevant inden for den medicinske forskning, til at vinde indpas i stadig flere områder af offentlig forvaltning og praksis.

Særligt betydningsfuld for evidensbegrebets vandring mod mere omfattende dele af den sociale og politiske praksis var Campbell, der under overskriften "the experimenting society" blev bannerfører for bevægelsen, der op gennem 80'erne og 90'erne fik vind i sejlene sammen med "new public management"-idealer for offentlig styringspraksis (Hammersley, 2013). Dette har bidraget til udviklingen af det såkaldte evalueringssamfund (Dahler-Larsen, 2009) med en institutionalisering af rutiner for (selv-) kontrol, standardisering, systematisering og dokumentation inden for hele den offentlige sektor (Hansen, 2016). Når professionerne møder krav om evidensbaserings, bunder det ofte i en rationalistisk input-output-tankegang, hvor snæver evidens tænkes at fungere som input til praksis med forventning om, at korrekt implementering vil lede til den ønskede effekt. Det dominerende evidensbegreb og den særstatus, begrebet tildeles som et angiveligt sikkert fundament for politiske beslutninger, kan ses som begrænsende for professionelles råderum, status og vidensarbejde.

Det professionelle skøn

Det professionelle skøn har traditionelt været væsentligt for forståelsen af professionel omgang med viden, og det har særlige normative implikationer for professionsudøvelsen.

Den professionelle praksis baserer sig på en kombination af vidensformer. Professionelle baserer sig på kodificeret, akademisk viden, "book knowledge", som Wackerhausen betegner det (2011, p. 13), og praktisk anvendelsesorienteret viden. Endelig, som en

tredje måde at arbejde med viden på, skal professionelle være i stand til intuitivt at agere på nye fænomener og aktivere den mest relevante tidligere indhentede viden til understøttelse af valget af handlemåder i situationen (Polanyi, 2012; Freidson, 2001; Grimen & Molander, 2008). Professionel praksis implicerer altså en mediering mellem forskellige vidensformer.

Grimen og Molander (2008) skelner mellem to elementer ved det professionelle skøn. På den ene side kan skønnet defineres som den kognitive ræsonnementsproces, der finder sted, når professionelle foretager denne vidensmæssige mediering på baggrund af ufuldstændige informationer i nye situationer. På den anden side kan skønnet betragtes som den professionelles negativt definerede frihed til at træffe beslutninger inden for et nærmere afgrænset råderum, der fx muliggør et valg mellem en række forskellige standardiserede metoder. Skønnet har afsæt i den professionelles dømmekraft (Pahuus & Eriksen, 2011).

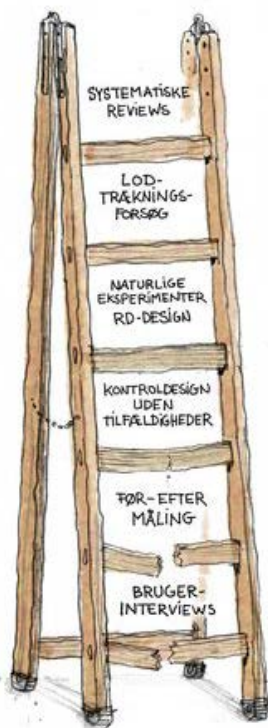
Det er et eksistentielt vilkår for det professionelle skøn, at der som udgangspunkt er usikkerhed, potentielt både i forhold til problemdefinition og problemløsning. Vidensgrundlaget for skønnet forhandles og udvikles, samtidig med at den professionelle praksis udøves. Usikkerhedselementerne ved skønnet giver en risiko for vilkårlighed eller bias (Grimen & Molander, 2008; Taylor, 2017). Der er således en normativ byrde forbundet med skønsudøvelse, hvor der meget let vil kunne opstå problemer med overholdelse af de tre almene, normative principper for professionel skønsudøvelse, som Grimen &

Molander (2008) identificerer, nemlig ligebehandlingsprincippet, reproducerbarhedsprincippet og individualiseringsprincippet. De to første principper adresserer ønsket om, at der skal være en vis komparativ konsistens imellem forskellige skøn, henholdsvis i forhold til ligebehandling af lige situationer og i forhold til ensartethed mellem forskellige skønsudøveres bedømmelser. Endelig anerkender individualiseringsprincippet, at enkeltindivider er forskellige og ofte har legitime krav på forskellig behandling, hvorfor et skøn, der tilgodeser individualiseringsprincippet, ofte vil kunne tilsidesætte et af de andre principper.

Skønnets normative byrde kan være tung at bære. Samtidig har offentligheden en legitim interesse i at minimere usikkerheden forbundet med professionel skønsudøvelse, der blandt andet afspejler sig i krav til åbenhed og dokumentation i den offentlige forvaltningspraksis. Professionelle skøn skal kunne underkastes efterfølgende stillen til regnskab, accountability. Det gør den dominerende evidensforståelses ideal om à priori at kunne identificere de beviseligt smarteste løsninger tillokkende.

Evidens og den gyldne standard

Den mest fremtrædende brug af evidensbegrebet indebærer en forestilling om, at nogle former for viden om, hvad der virker, er bedre end andre. Forskellige metoder til at opnå viden vurderes at have forskellig kvalitet (Rieper & Hansen, 2007). Dette afspejles fx ved illustrationen, der tidligere blev brugt af Trygfondens Børneforskningscenter ved Aarhus Universitet, figur 1.



Figur 1. Illustration fra Trygfondens Børneforskningscenters hjemmeside <http://childresearch.au.dk/metoder/> accessed 27. april 2017.

Inden for en sådan snæver fortolkning af evidensbegrebet er der således opstillet et hierarki over, hvilke metoder der vurderes som mest valide for et forskningsdesigns evne til at tilvejebringe evidens, der altså her forstås som sikker, beviselig viden, der ikke kan betvivles (Krogstrup, 2011).

Rangordningen af evidensgrundlagets validitet og kvalitet sker efter den evidenshierarkiske logik udelukkende på baggrund af metodisk design. I toppen af hierarkiet ligger systematiske reviews, der omfatter forskningsoversigter og metaanalyse. Det er en metode,

der forsøger at sammenfatte en række tidligere studier af en specifik indsats for at kunne opskalere den samlede viden om indsatsens effekt. Forud for disse metastudier foregår en rangordning af de primærstudier, der indgår i dem, på baggrund af en validitetsvurdering af de anvendte metoder. Her gælder det ifølge evidenshierarkiet, at viden om virkningen af en indsats i så stor udstrækning som muligt skal være baseret på randomiserede, kontrollerede forsøg (RCT), der udgør den "gyldne standard" (Krogstrup, 2011; Rieper & Hansen, 2007). En RCT kan finde sted, når undersøgelsens subjekter, hvad enten der er tale om mennesker som ved et medicinsk eksperiment, organisatoriske enheder eller noget helt andet, tilfældigt inddeles i mindst to grupper. Den ene gruppe tildeles indsatsen, hvis effekt ønskes undersøgt, hvorimod en anden gruppe skal fungere som kontrolgruppe ved at få fx placebo eller den almindeligt udbudte indsats. Efter implementering af indsatsen måles effekterne i både kontrol- og indsatsgruppe. Kontrolgruppen fungerer så som realisering af, hvad der ville være sket med indsatsgruppen, hvis den ikke var blevet tilbudt indsatsen (John, 2016).

Hvis der ikke er brugt randomisering, må man tage til takke med viden baseret på fx kvasi-eksperimenter og naturlige eksperimenter, hvor man må forlade sig på, at forskeren har styr på, hvilke variable der bør kontrolleres for. Længere nede af stigen kommer procesevalueringer, kvalitative metoder, ekspertvurderinger, brugerundersøgelser etc.

Den evidenshierarkiske logik er et udtryk for kausalitetstænkning. At forfølge en snæver form for evidens indebærer et

ønske om at kunne etablere deterministisk kausalitet. Når effektevaluatorene hævder at have fundet evidens for fx en indsats, så mener de at have isoleret en kausal sammenhæng, der kan bruges som bevisbyrde for, at den målte effekt kan tilskrives den anvendte indsats. Effekten er indtruffet, som et direkte resultat af, at der blev brugt en bestemt indsats (Krogstrup, 2011). Den evidenshierarkiske tankegang indebærer således et positivistisk ideal om afdækning af sande årsagsforklaringer.

Det er netop evnen til at kunne isolere effektmålet til udelukkende at måle på nettosammenhængen mellem indsats og effekt gennem den tilfældige tildeling af subjekter til kontrol- og interventionsgruppe, der er en af de væsentligste styrker ved RCT. Modsat de statistiske metoder, der fx tages i anvendelse ved kvasi-eksperimenter, kan RCT-studier også tage højde for uerkendte faktorer, der eventuelt måtte spille ind, da randomiseringen – i al fald inden for rammerne af hvad der vil være statistisk tilfældigt med et eksperiment af en given størrelse – vil eliminere alle biases i selektionen af de to grupper. Dette giver en høj intern validitet af et RCT forskningsdesign (Rieper & Hansen, 2007).

Kritik af evidensbegrebet

Det er svært at være imod ideen om, at beslutningsprocesser skal basere sig på viden om, hvad der virker bedst. Imidlertid knytter der sig en række problemer til at træffe beslutninger på baggrund af denne snævre form for evidens.

1) Evidensbegrebets

indbyggede uantastelighed

På et metaplan kan evidensbegrebet kritiseres for sin entydigt positive, værdis-

Det er et eksistentielt vilkår for det professionelle skøn, at der som udgangspunkt er usikkerhed, potentielt både i forhold til problemdefinition og problemløsning.

Vidensgrundlaget for skønnet forhandles og udvikles, samtidig med at den professionelle praksis udøves.



mæssige bias, der medvirker til at bringe kritiske røster til tavshed. Begrebets "rhetorical effect is to discredit opposition. After all, who would argue that practice should not be based on evidence" (Hammersley, 2013, p. 15). En udvidelse af evidensbegrebet, som denne artikel argumenterer for, indebærer en risiko for, at evidensbegrebet bliver en "tom betegnelse" (Laclau & Mouffe, 2002). Det vil sige, at begrebet bliver for løst defineret til, at dets betydning entydigt kan afgøres. Imidlertid er udfyldningen af mening i begrebet kontingent og en nærmere identifikation af alternativer² væsentlig for forståelsen af den mere politiserede brug af begrebet.

2) RCT-baseret evidens giver for smalt vidensgrundlag

For at kunne bruge den indsamlede

viden til noget, er det ikke nok at kunne måle effekten af en indsats. Det er ofte lige så væsentligt at vide, hvorfor en indsats virker eller ikke virker, at kunne åbne den kausale sammenhængs sorte boks og undersøge, hvad der sker i det, der ofte omtales som en "kæde" mellem implementering og målt effekt. Samtidig giver eksperimentelle designs ikke nogen viden om involverede aktørers oplevelser af processen og indsatsen. Og endelig er der kun sjældent mulighed for at indarbejde en hensyntagen til forskellige indsatsers relative omkostningseffektivitet. Når nu realiteten er den, at vi ikke har ubegrænsede ressourcer, så vi vil også godt have mulighed for at vurdere, hvor omkostningseffektive forskellige indsatser er i forhold til hinanden (Rieper & Hansen, 2007).

3) Praktisk-teknisk realiserbarhed

En type kritik af imperativet om evidensbaseret af den offentlige sektors aktiviteter peger på, at det i praksis er så godt som umuligt at opstille RCT-eksperimenter, der kan tilvejebringe evidens. I den virkelige verden, hvor professionel praksis udspiller sig, er det ualmindelig svært – hvis ikke umuligt – at "holde alt andet lige", som metodeforskrifterne tilsiger for at kunne tilskrive effektmålet den indsats, man ønsker at vurdere. Hvis man er interesseret i at fastslå, om der er en deterministisk, kausal relation mellem en indsats og dens effekt – altså om den virker, og om de observerede effekter nu også med sikkerhed er forårsagede af den – så ligger der en stor metodisk udfordring i at isolere de variable, som man vil undersøge (Krogstrup, 2011, p. 108ff).

Det er svært at være imod ideen om, at beslutningsprocesser skal basere sig på viden om, hvad der virker bedst. **Imidlertid knytter der sig en række problemer til at træffe beslutninger på baggrund af denne snævre form for evidens.**



I særlige tilfælde vil man meningsfuldt kunne opstille randomiserede felteksperimenter, naturlige eksperimenter eller kvasi-eksperimenter (se fx Andersen et al., 2012) og argumentere for, at resultaterne fundet her ligger tættere på "det naturlige". Der kan desuden være en række etiske problemer ved at opstille eksperimenter (Rieper & Hansen, 2007). Den praktisk-tekniske kritik af evidensbegrebet godtager ofte den ideale forestilling om et epistemologisk videnshierarki, men påpeger, at det er uhyre vanskeligt at operationalisere og implementere arbejdet med evidensskabelse inden for mere praksisnære og dynamiske områder af samfundet, der ikke udspiller sig i et laboratorium.

4) Overførbarhedsproblemet

Netop RCT'ernes laboratorieagtige logik er afsæt for et andet væsentligt kritikpunkt fremført af Cartwright og Hardie

(2012), der fokuserer på problemet med at kunne overføre RCT-baseret evidens til en anden kontekst. For at kunne basere ny social eller politisk aktivitet på evidens fundet i tidligere studier, må man antage, at konklusionen om, hvad der virkede i den målte situation, meningsfuldt kan bruges som bevis på, at det også vil virke i en anden situation et andet sted. Og dette er en højst skrøbelig antagelse, vurderer de. "RCTs ... only secure the starting point in the long road from 'it works there' to 'it will work here'..." (Cartwright & Hardie, 2012, p. 123). De randomiserede, kontrollerede forsøg har altså ofte problemer med den eksterne validitet, dvs. med at kunne overføre resultaterne til en større, mere generel kontekst (Rieper & Hansen, 2007). At flytte en testet indsats i tid og sted kan implicere en lang række anderledes kontekstuelle forhold, og for at kunne gøre antagelsen om, at indsatsen virker

i andre sammenhænge, må man forsøge at forstå, hvad der får den til at virke (Cartwright & Hardie, 2012).

5) Anvendelsesproblemet

Bag evidensbevægelsen ligger en antagelse om, at tilvejebringelsen af tilstrækkelig valid viden i sig selv udgør argumentet, der trumfer alle andre, i beslutningsprocesser om fremtidige politiske og sociale indsatser. Det forholder sig imidlertid ikke helt sådan. Inden for evalueringsforskning har det længe været et velkendt fænomen, at der ofte er en uoverensstemmelse mellem en række policy evaluerings udtalte målsætning om at vurdere effekter med henblik på at kvalificere fremtidig beslutningstagning, og så den måde hvorpå evalueringsresultaterne finder anvendelse i beslutningstagningen (Albæk, 1988; Dahler-Larsen, 2004; Shulock, 1999; Larsen & Lassen, 2001).

Selvom vi forestillede os, at det rent faktisk var muligt at forudsige, hvad der ville være den bedste indsats med henblik på at afhjælpe et givet samfundsmæssigt problem (se afsnittet om ontologisk kritik nedenfor for en problematisering af selve problemdefinitionen), så er det bestemt ikke sikkert, at denne viden i sig selv ville være tilstrækkelig til, at beslutningen om at implementere indsatsen ville blive truffet. Det skyldes ifølge en række politologer, at den politiske beslutningsproces ikke kan betragtes som entydig, lineært fremadskridende og rationel (Larsen & Lassen, 2001; Albæk, 1988). Som Schwandt (2015) påpeger, er anvendelse af viden nok nærmere en kollektiv, interaktiv samskabelsesproces, hvor konfliktende hensyn, alternative dagsordener og dilemmaer skal afvejes og forhandles i specifikke kontekster. Hvor idealforestillingen om evidens tilsiger, at evidens skal danne grundlag for beslutningsprocesser, er praksis måske helt modsat. Produktionen af evidens kan lige så vel være baseret på policy med henblik på at legitimere beslutningsprocesser, som Pawson (2006, p. viii) samt Stoker og Evans (2016, p. 16) er inde på. Ingen af dem afskriver dog evidensbegrebet. Pawson (2006) arbejder for mere systematisering og større inklusivitet i det metodiske arbejde med at indsamle evidens og opbygge teori, samtidig med at han henleder opmærksomheden til væsentligheden af at forstå, *hvordan* indsatser virker (mekanismen), fremfor hvor godt de virker. Stoker og Evans (2016) advokerer for metodisk pluralisme og for involvering af bruger-/borgerperspektiver i en samproducerende, åben og udviklende forståelse af, hvilke vidensformer forskningen bør arbejde på at tilvejebringe samfundet.

6) Ontologisk kritik

– *virkelighedens status*

De problemer, der til dagligt varetages af fx velfærdsprofessionelle, er ikke altid ligetil at putte i kasser med henblik på kategorisering og sammenligning. Alt for mange situationsbestemte variable interagerer, og både strukturer og aktører udvikler sig kontinuerligt. Ud over at præsentere et praktisk problem ved forsøg på at måle effekter af indsatser, kan kompleksitet opfattes som et selvstændigt ontologisk vilkår, som udelukker en snæver evidensforståelse. Problemer og beslutninger i de professionelles praksis er komplekse og 'wicked' i modsætning til tamme og simple (Rittel & Webber, 1973; Krogstrup, 2011; Patton, 2011; Hammersley, 2013). Samfundsvidenskabelig kompleksitetsteori påpeger desuden betydningen af emergens, unik responsivitet og innovation. Det vil ofte være hensigtsmæssigt at tænke nyt og prøve sig frem under responsiv refleksion. Endemål kan være ukendte eller opstå hen ad vejen, og der er ikke altid konsensus om hverken mål eller midler. Virkeligheden udvikles i kraft af en vedvarende reaktion på nye situationer (Patton 2011). For overhovedet at kunne blive enige om at måle på en given indsats, er vi nødt til allerede at have identificeret det problem, vi ønsker at afhjælpe med indsatsen. Dette kan i sig selv være ualmindelig vanskeligt. Når der fx fra politisk hold bliver meldt ud, at Danmark skal have "verdens bedste folkeskole" (se også Krogstrup, 2011, p. 83), hvad er det så lige, der menes? Og kan politikere, lærere, evaluører, uddannelsesforskere, kommende arbejdsgivere, elever og forældre – for bare at nævne de mest iøjefaldende interessenter – nogensinde opnå en

fælles forståelse af, hvad en god folkeskole er? Når de opstillede politiske mål er uklart definerede, bliver forsøget på at indhente evidens meget let afkoblet fra den virkelighed, som forskellige parter oplever som meningsfuld. En væsentlig del af den ontologiske kritik af evidensbevægelsen er baseret på, at den snævre evidensforståelse er ude af harmoni med og dårlig til at rumme centrale aktørers oplevelser og erfaringer (Biesta, 2007; Biesta, 2010; Moos et al., 2005).

7) Epistemologisk kritik

– *vidensbegrebets status*

En grundlæggende anke mod det snævre evidensbegreb angår vidensbegrebets status og har således epistemologiske implikationer. Forestillingen om, at RCT-baseret viden er den mest valide form for viden er alt for forsimplet og helt ude af trit med virkeligheden. Selv den mest strenge eksperimentelt frembragte form for data er konstrueret viden, hævder kritikerne (Schwandt, 2015; Stone, 2016). Ifølge denne evidenskritiske position findes der ingen kontekstfri, neutral viden. "Viden er ikke entydig, hvilket vil sige, at den kan udfordres socialt" (Krogstrup, 2011, p. 86.) Det, der fremlægges som evidens, vil til enhver tid være en fortolkning af data. Selv tilsyneladende entydige, deskriptive karakteristik, som blot kvantificerer, kan betragtes som kontingente, sociale konstruktioner. De mentale processer, der ligger bag kvantificering og kategorisering er hverken naturgivne eller uanfægtelige (Stone, 2016).

"Those who count have to know – or rather they must decide – what counts as an instance of the thing they are

counting. Counting always involves deliberate decisions about categorizing" (Stone, 2016, p. 160, original fremhævelse).

Evidenskritik, der anfægter de positivistiske idealer om at afdække "sande kausalsammenhænge", hævder at praksis ikke lader sig observere og analysere ved hjælp af et distanceret, "objektivt" forskningsblik. Forskellige positioner foreslår i stedet at arbejde med konfigurativ kausalitet eller plausibel/probabilistisk kausalitet, der betoner henholdsvis komplekse, gensidige interaktionsmekanismer og sandsynliggørelse af indvirkninger (Rieper & Hansen, 2007; Krogstrup, 2016; Pawson, 2006; Dahler-Larsen, 2004). Den epistemologiske kritik af den snævre, RCT-baserede evidens påpeger, at alle former for vidensarbejde indebærer mangler og partielle blik. Det perfekte design findes ikke, og springene mellem problemidentifikation, opstilling af målemetode, indsamling af data og formidling af resultater indebærer alle fortolkning og vurdering (Hammersley, 2013).

Mod et inklusivt evidensbegreb

I denne artikel er det hensigten at udvide evidensbegrebet til at omfatte andre vidensformer end dem, der baserer sig på eksperimentel metode. Det er en pointe her, at ingen form for viden kan hævde at være universelt "ren" og relevant. Metodisk stringens og høj intern validitet er ikke mere væsentlige parametre for god evidens end situationsbestemt meningsfuldhed.

Selvom evidensbegrebet som anført er højst omdiskuteret fastholdes termen, for at understrege, at kritikpunkterne er væsentlige og må adresseres. Fremfø-

ringen af det skønnede evidensbegreb er således et bidrag til at holde debatten om evidensbegrebet åben, så den snævre RCT-funderede evidensforståelse ikke står alene om at fylde mening i det udflydende evidensbegreb.

Som illustreret i Figur 2 skal den skønnede evidens forstås som et inklusivt evidensbegreb situeret i den professionelle praksis. Det er tanken, at der her er en cirkulær kontinuerlig udvikling af både den professionelle praksisudøvelse og af en evidens, der omfatter mange former for akkumuleret viden i supplement til hinanden. Skønnet evidens er både den viden, professionelle baserer sig på og agerer ud fra, ligesom det er viden, de bidrager til.

Den epistemologiske kritik af det snævre evidensbegreb indebærer, at al viden kan underkastes kritik. Alle vidensformer er partielle og for størst mulig brugbarhed af vidensgrundlaget

må det være inklusivt og mangesidet. Derfor indeholder den skønnede evidens også en kontinuerlig praktisk evidensanvendelse under responsiv refleksion med henblik på udvikling af evidens. Selve diskussionen og den selvrefleksion, den kan afstedkomme i vidensarbejdet, er væsentlig for kvaliteten af evidensgrundlaget.

Som evidenskritikken ovenfor indikerer, er en mekanisk overførsel og implementering af metoder, der har fået stemplet "evidensbaserede", fra en kontekst til en anden problematisk (se også Reime, 2016). Møller (2017a og 2017b) eksemplificerer, at der nu inden for det socialfaglige område i stedet er tendenser til, at brugen af evidensbaserede metoder kan finde sted under inddragelse af lokal, praksis-baseret erfaring og situeret viden hos de professionelle aktører. Når enhver form for viden (inklusive evidens) om, hvad der synes at virke godt, skal bringes i anvendelse, vil der skulle ske



Figur 2. Den skønnede evidens

en translationsproces fra policy niveau til praksis (Møller 2017a og 2017b; Reime 2016). Det er professionerne, der har fingeren på pulsen i forhold til den konkrete kontekst. Deres særlige, faglige fokus og erfaringsbaserede viden giver dem en væsentlig nøglerolle i forhold til evidensbaseret praksis. Evidensbaseret praksis betyder dog ikke, at der skal udvikles ét stabilt bidrag til den bedste udformning af praksis. Udviklingen af praksis er en kontinuerlig proces, som professionerne spiller en privilegeret rolle i, og hvor der kontinuerligt skal passes til og justeres efter den aktuelle udvikling. Der vil skulle ske en situeret adaptering på baggrund af et professionelt skøn. Det professionelle arbejde med evidens er dobbelt. Professionerne skal både forstås som aktive medskabere af en situeret "bedste praksis", og som aktive bidragsydere til udviklingen af en mere inklusiv evidens til gavn for både nye generationer af professionelle og for den fremtidige understøttelse af praksis. Udvikling af en inklusiv evidens stiller krav til professionerne om at italesætte og dele erfaringer og "tavs" viden, og om at foretage systematisk, selvrefleksiv observation og evaluering, der kan udgøre fremadrettede bidrag til evidensbasen. Selvom illustrationen af det skønnede evidensbegreb i figur 2 udelukkende viser bidrag fra den professionelle praksis til evidens, vil der selvfølgelig kunne være en række andre kilder til dette inklusive evidensbegreb.

Phronesis – et imperativ for evidensarbejde i professionel praksis

Det dobbelte, professionelle arbejde med at bringe evidens i meningsfuld, situeret anvendelse og producere nye bidrag til evidens er at betragte som phronetisk. Herhjemme har særligt

Flyvbjerg (2001) arbejdet på at udvikle samfundsvidenskabelig forskning i retning af en mere praksisnær retning, hvor der tages afstand fra det uopnåelige ideal om sand, objektiv viden efter naturvidenskabeligt forbillede. Med henvisning til Aristoteles' phronesis-begreb opstiller Flyvbjerg et epistemologisk korrektiv til den dominerende tilgang.

Den dominerende forestilling om en universel, objektiv og kontekstuaafhængig videnskab, der opererer med afdækning af "sandhed" og som RCT-evidensbegrebet er affødt af, trækker tråde tilbage til Aristoteles' begreb om episteme, som en af en række intellektuelle dyder³. Den måde at forholde sig til viden på, som episteme-begrebet betegner, har fået lov at dominere hele vores forestilling om, hvad videnskab er, hvorved øvrige dyder er blevet inferiøre. I mange engelske oversættelser bliver episteme sågar betegnet som "science".

Techné er en form for dyd, der knytter sig til praktisk skabende arbejde fx i forbindelse med håndværk og kunst. Techné er kontekstafhængig og baseret på konkret erfaring med en foranderlig praksis. Som techné er phronesis også kontekstsensitiv og varierende efter omstændighederne. I modsætning til episteme, der i høj grad er en instrumentel rationalitet, er phronesis værdiorienteret og orienteret mod det etiske, hvilket indebærer deliberation og dømmekraft. Phronesis er som techné relateret til praksis, men hvor techné er specifik i forhold til konkret know how i produktionsorienteret forstand, har phronesis fokus på erfaringer og handling. Professionerne er særligt privilegerede videnskudpunkter, idet de netop har den største erfaringsbaserede viden om

praksis, som de er vant til at foretage professionelle skøn omkring.

Dette gør phronesis særligt relevant for professionelles vidensarbejde og den skønnede evidens. Flyvbjerg mener da også, at den menneskelige læring er en trinvis udvikling fra et novice-stadie til ekspertstadiet, der opnås gennem erfaringsbaseret læring i praksis, og som indebærer evnen til phronesis. Phronesis kan forstås som en slags etisk-praktisk sans, der er retningsgiver for konkrete handlinger på baggrund af en værdianalyse, der betænker "det gode". Phronesis bruges altså til at betegne menneskets aktive overvejelse omkring retningen på den ønskede udvikling og karakteren af det gode liv. Phronesis-begrebets orientering mod etik og "det gode" kan give anledning til den opfattelse, at udelukkende gode gerninger – i den udstrækning sådanne meningsfuldt vil kunne identificeres og isoleres – kan henhøre til phronesis. Men det er ikke udfaldet af en tredjeparts epistemiske sandhedsdom over graden af godhed, der afgør om en intellektuel aktivitet tilhører det phronetiske område. Det er selve den aktivitet, der består i at undersøge, hvorvidt noget vil være godt eller skidt i henhold til specifikke værdier og interesser, der er phronetisk. (Flyvbjerg, 2001, p. 57)

I Flyvbjergs brug af begrebet som rettesnor for socialvidenskaberne indtænkes desuden magtbegrebet, der ikke er immanent i det Aristoteliske phronesis-begreb. Han er her inspireret af Foucaults fokus på relationel magt, der er væsentligst at undersøge i interaktionen mellem reelt handlende mennesker eller agerende institutioner. Den phronetiske forskningsmæssige

position overtager også den normative, emancipatoriske opfattelse af videnskabens rolle i samfundet, som Flyvbjerg (2001, p. 128) identificerer hos Foucault. Uden at underkende værdien af magtperspektiver, der forsøger at lokalisere magtcentre og afdække magten som en (ofte ujævnt fordelt) ressource eller besiddelse, finder Flyvbjerg det således mere frugtbart at indhente viden om, hvordan magt i praksis udspiller sig mellem aktører.

Flyvbjergs brug af – og bidrag til – teorier af den form, som han ville kategorisere som generaliserende, epistemisk videnskab, gør Flyvbjerg sårbar over for sin egen kritik af sådanne rationelle, videnskabelige forehavender (Thiele, 2002).

One judges well by discerning how the comprehensive informs the contingent, how the abstract informs the concrete, and vice versa. Because judgment always pertains to context, it cannot be reduced to a wholly deductive, analytic exercise. Yet, good judgment is not wholly particularistic. It is a matter of understanding when, where, and how universals are applicable, how they must be modified in particular circumstances, and when they cease to apply altogether [...] After all, it is only by means of a thorough knowledge of rules that one learns where, when, and how to break them (Thiele 2002, p. 276).

Hvor Flyvbjerg insisterer på, at ekspertens intuitivt handlingsrettede phronesis unddrager sig italesættelse og forbliver "tavs", er det imidlertid hensigten her, at gøre det til et normativt imperativ for professionelt

arbejde at forsøge at sætte ord på. Det skønnede evidensbegreb kombinerer en række forskellige kilder til viden og indfører samtidig en målsætning om at bidrage til yderligere skønnet evidens gennem systematisk og dokumenteret selvrefleksion.

Konklusion

På baggrund af en række problemer relateret til den dominerende, snævre evidensforståelse baseret på eksperimentel metode har denne artikel fremsat et inklusivt korrektiv, nemlig den skønnede evidens. Flyvbjergs phronesis-forståelse byder sig til som et begrebsapparat, der kan understøtte en praksisnær, situeret omgang med evidens i det professionelle arbejde. Opgøret med episteme som ideal for vidensproduktion er helt parallelt med kritikken af et evidensbegreb, der idealiserer eksperimentelle metoder.

Udviklingen af et inklusivt, phronetisk evidensbegreb, den skønnede evidens, søger at bygge bro over videnskabs-teoretiske grøfter og overvinde teknokratisk rationalisme ved at inkludere kontekst og værdier. Under inddragelse af det professionelle skøn er det situationen i den konkrete professionelle praksis, der afføder den for øjeblikket bedste vurdering.

Phronetisk professionsudøvelse er dog ikke frisat for ansvar i en relativistisk forstand, hvor en indsats kan være lige så god som en anden. Derimod ligger der en fordring om hele tiden at gøre skønsgrundlaget til genstand for undersøgelse, refleksion, argumentation og udvidelse. Evidensgrundlaget er varieret og mangeartet og indeholder både den erfaringsbaserede viden, der

ofte betragtes som tavs, samt mere klassisk-videnskabelige, episteme-inspirerede bidrag til en kumulativ viden. Det fornemste vidensarbejde ligger i den reflektivt phronetiske håndtering af en inklusiv evidens, hvor både validitet og værdi er i fokus.

Selvom Møller (2017a, p. 225) finder, at "arbejdet med systematisk feedback i den grad taler til de fagprofessionelles indre drivkraft og motivation", så bærer den skønnede evidensform skønnets byrde med sig, ligesom magtrelationer vil kunne spille ind. Evidens er ikke uantastelig, og der er risiko for variation, bias og uenighed (Grimen & Molander, 2008), som dynamisk og systematisk italesættelse skal omsætte til bedre praksis og bedre evidens. – Måske endda med potentiale for øget rygstød og legitimering af det professionelle arbejde (Buus & Rasmussen, 2015).

Fremadrettet kan der ligge et potentiale for professionsforskningen og for de professionelle i at høste viden, der er baseret på faglige skøn, og omsætte den, så den kan "help inform practical reason or phronesis, that is, the ability to make intelligent decisions in particular circumstances" (Caterino & Schram, 2006, p. 1). ♦

ACKNOWLEDGEMENT

Tak til to anonyme reviewere for ualmindeligt brugbar feedback på første udkast til artiklen.

REFERENCES

- Albæk, E. (1988). *Policy Evaluation: Design and Utilization*. Conference paper, Aarhus Universitet, 1-24.
- Andersen, S.C. & Jakobsen M. (2012). *Hovedresultater fra en eksperimentel undersøgelse af en forstærket indsats over for førskolebørn med dansk som andetsprog*. Institut for Statskundskab, Aarhus Universitet
- Biesta, G. (2007). Why "What Works" Won't Work: Evidence-Based Practice and the Democratic Deficit in Educational Research, *Educational Theory*, 57,1, 1-22.
- Biesta, G. (2010). *Good Education in an Age of Measurement*. Paradigm Publishers.
- Buus, A.M. & Rasmussen, P. (2015). Når evidensbaserede metoder møder den pædagogiske hverdag. In T. Næsby (ed.) *Evidens i pædagogens praksis*. Frederikshavn: Dafolo.
- Cartwright, N. & Hardie, J. (2012). *Evidence-Based Policy: A Practical Guide to Doing It Better*. Oxford: Oxford University Press.
- Caterino, B. & Schram, S.F. (2006). Introduction. In S.F. Schram & B. Caterino (eds.) *Making Political Science Matter*, New York University Press.
- Dahler-Larsen, P. (2009). *Den rituelle refleksion*, Viborg: Syddansk Universitetsforlag.
- Dahler-Larsen, P. (2004). *Evaluerings og magt*, Magtudredningen.
- Flyvbjerg, B. (2001). *Making Social Science Matter*, Cambridge: Cambridge University Press.
- Freidson, E. (2001): *Professionalism. The Third Logic*, Cambridge: Polity Press.
- Grimen, H. & Molander, A. (2008): "Profesjon og skjønn". In A.Molander & L.I. Terum (eds.), *Profesjonsstudier* (pp. 179-215). Oslo: Universitetsforlaget.
- Hammersley, M. 2013. *The Myth of Research-Based Policy and Practice*, Los Angeles: Sage.
- Hansen, H.F. (2016). Evalueringsstyring. In T. Bredgaard (ed.) *Evaluerings af offentlig politik og administration* (pp. 95-114). København: Hans Reitzels forlag.
- John, P. (2016). Randomised controlled trials. In G. Stoker & M. Evans (eds.) *Evidence-based policy making in the social sciences* (pp. 69-82). Bristol: Policy Press.
- Krogstrup, H.K. (2016): *Evalueringsmodeller* (3. edition). Hans Reitzels Forlag.
- Krogstrup, H.K. (2011). *Kampen om evidens*, København: Hans Reitzels Forlag.
- Laclau, E. & Mouffe, C. (2002). Hinsides det sociale positivitet. In C. Jensen & A.D. Hansen (eds.) *Det radikale demokrati – diskursteoriens politiske perspektiv* (pp. 37-100). Roskilde Universitetsforlag.
- Larsen, F. & Lassen, M. (2001). Evalueringen set i et policy-perspektiv. In P. Dahler-Larsen & H. K. Krogstrup (eds.), *Tendenser i evaluering* (pp. 46-60). Odense: Syddansk Universitetsforlag.
- Moos, L. et al. (2005). *Evidens i uddannelse?* København: Danmarks Pædagogiske Universitets Forlag.
- Møller, A.M. (2017a). Evidens i praksis. In E.M. Ghin et al. (eds.) *Styring og evaluering i den offentlige sektor* (pp. 209-234). København: Hans Reitzels Forlag.

Møller, A.M. (2017b). Evidensdagsordenens metamorfoser i dansk socialpolitik. *Politik*, 2:20, 68-87.

Pahuus, A.M. & Eriksen, C. (2011): "Hvad er dømmekraft? – den levende arv fra Aristoteles, Kant og Løgstrup". In M.B. Johansen, & S.G. Olesen, (eds.), *Professionernes sociologi og vidensgrundlag* (pp. 43-65). Viborg: ViaSysteme.

Patton, M.Q. (2011). *Developmental Evaluation*. New York: The Guilford Press.

Pawson, R. (2006). *Evidence-based policy. A realist perspective*, London: Sage.

Polanyi, M. (2012). *Den tavse dimension*, (orig. *The tacit Dimension*) Forlaget Mindspace [1966].

Reime, M.A. (2016). Discourses in Residential Childcare and Possibilities for Evidence-Based Practice. *Professions & Professionalism*, 6:1, 1-15.

Rieper, O. & Hansen, H.F. (2007). *Metodedebatten om evidens*. København: AKF Forlaget.

Rittel and Webber (1973). Dilemmas in a General Theory of Planning. *Policy Sciences*, 4, 155-169.

Schwandt, T.A. (2015). Iagttagelser af anvendelse af videnskabelig evidens i politik og praksis. *Økonomi & Politik*, 1, 88, 5-12.

Shulock, N. 1999. The Paradox of Policy Analysis: If It Is Not Used, Why Do We Produce So Much of It? *Journal of Policy Analysis and Management*, 18:2, 226-244.

Stoker, G. & Evans, M. (2016). *Evidence-based policy making in the social sciences*. Bristol: Policy Press.

Stone, D. (2016). Quantitative Analysis as Narrative. In M. Bevir & R. A. W. Rhodes (ed.), *Routledge Handbook of Interpretive Political Science* (pp. 157-170). London: Routledge.

Taylor, B.J. (2017). Heuristics in Professional Judgment: A Psycho-Social Rationality Model, *British Journal of Social Work*, 47: 1043-1060.

Thiele, L.P. (2002) Review. *The Journal of Politics*, 64, 1, 274-276.

Wackerhausen, S. (2011). Professionsidentitet, sædvane og akademiske dyder. In N.B. Hansen & J. Gleerup (eds.), *Videnteori, Professionsuddannelse og Professionsforskning* (pp. 13-30). Syddansk Universitetsforlag.

ENDNOTES

¹ I det følgende bruges termen indsats som samlebetegnelse for de aktiviteter, interventioner mm., som iværksættes med henblik på opnåelsen af overordnede, samfundsmæssige mål ved brug af dertil allokerede ressourcer.

² Rieper og Hansen (2007) foreslår at arbejde med sidestillede evidensstypologier som alternativ til en rangordning i et generelt evidenshierarki.

³ Aristoteles' to øvrige dyder fra *Den Nikomacheiske Etik*, sophia og nous, gennemgås ikke nærmere her, da de ikke har betydning for argumentet i denne fremstilling.